**1. Analytical Approach**

**a. Data Sources to Analyze:**

* **Network Logs:** Identify abnormal outbound connections, unauthorized data transfers, VPN usage anomalies, login times outside of normal hours.
* **Employee Access Records:** Look for unusual file access patterns, privilege escalation, or attempts to access restricted data.
* **Email Communications:** Apply NLP techniques to detect insider threat signals (e.g., sentiment analysis, keyword detection related to exfiltration or dissatisfaction).
* **System Event Logs:** Look for evidence of data staging, USB usage, or tampering with logs.

**2. Challenges & Solutions**

| **Challenge** | **Approach** |
| --- | --- |
| **Legitimate vs Malicious Behavior** | Use baseline profiling: compare individual behavior against peer groups and historical norms. Apply anomaly detection algorithms (e.g., Isolation Forest, Autoencoders). |
| **Privacy Concerns** | Minimize intrusion by anonymizing data during initial exploration. Only deanonymize when a strong case emerges. Secure and audit access to sensitive data. |
| **Data Volume and Complexity** | Use distributed processing tools like Apache Spark. Leverage ELK stack for log aggregation and visualization. |
| **False Positives** | Integrate feedback from HR/security analysts and continually refine detection models. Incorporate contextual data (e.g., project deadlines, approved remote access). |

**3. Ethical and Transparency Considerations**

* **Privacy Rights:** Ensure analysis adheres to corporate policy, GDPR/CCPA guidelines, and data minimization principles.
* **Transparency:** Communicate investigation scope and purpose clearly to employees. Log and justify all data access and actions.
* **Stakeholder Communication:**
  + **Technical Audience:** Provide anomaly detection metrics, visualizations, and feature importance.
  + **Non-Technical Audience:** Use analogies, timelines, and simple charts (e.g., the traffic chart like the one shown) to tell the story of potential malicious behavior.

**4. Strategies for Ethical & Effective Communication**

* **Use dashboards** (e.g., Power BI, Tableau) with filters by user/department.
* **Write clear summaries** separating facts from inferences.
* **Involve HR/legal teams** early when human behavior is under investigation.

2. GENERATE the visualization below in data Science. Add all relevant screen shots as well from your program. Also share the URL of your GITHUB (Where you have uploaded your work) so that I can simulate the same.



